HEMIPTERAN POLO-LIKE KINASES

Field of the Invention

The present invention relates to compositions that are useful in agrochemical, veterinary or pharmaceutical fields. In particular, the invention relates to nucleotide sequences that encode polypeptides that are useful in the identification or development of compounds that affect polo-like kinase activity, indicating that such compounds may be useful as pesticides or as pharmaceuticals.

Background of the Invention

Polo-like kinase ("PLK") proteins play an important role in regulating the cell cycle. Namely, it is suspected that PLK proteins are necessary for the maturation of mitotic centrosomes and the initiation of mitosis and participate in cytokinesis (Chase et al., genesis, 26, pp. 26-41 (2000); Donaldson et al., J. of Cell Biology, Vol. 153, No. 4, pp. 663-675 (2001); Hamanaka et al., J. of Biological Chemistry, Vol. 270, No. 36, pp. 21086-21091 (1995); Lee et al., Proc. Natl. Acad. Sci., Vol. 95, pp.9301-9306 (1998); Carmo Avides et al., Nature Cell Biology, Vol. 3, pp. 421-424 (2001); Elsa Logarinho and Claudio E. Sunkel, J. of Cell Sci., 111, pp. 2897-2909 (1998); Charles et al., Current Biology, Vol. 8, No. 9, pp. 497-507 (1998); Ouyang et al., J. of Biological Chemistry, Vol. 272, No. 45, pp. 28646-28651 (1995); and Qian et al., Mol. Cell Biol., Vol. 19, No. 12, pp. 8625-8632 (1999)). As such, there is a desire to develop ways to target these proteins as a means of identifying biologically active compounds (Anderson et al., WO 99/09146, published February 25, 1999 and Erikson et al., U.S. Patent 6,358,738 B1, issued March 19, 2002).

Polo-like kinase proteins have been expressed and cloned from various eukaryotic organisms, for example, mammals, *Xenopus laevis*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae* (Chase et al., Donaldson et al., Lee et al., Elsa Logarinho and Claudio E. Sunkel, Charles et al., Qian et al., and U.S. Patent 6,358,738). However, most of these proteins tend to involve humans, rabbits and mice or simple invertebrate organisms, rather than more complex invertebrate organisms, such as those from

the order of *Hemiptera*, including, but not limited to, the organism *Aphis gossypii*. As a result, there is a need for PLK proteins that are non-human, involve more complex invertebrate species, such as invertebrate organisms from the order of *Hemiptera*, more particular the organism *Aphis gossypii*, and that exhibit activity as serine and threonine kinases.

Methods for measuring the ability of a compound to affect PLK activity are known to those skilled in the art. For example, one of ordinary skill in the art would know that growth inhibition measurements in oocytes expressing a functional PLK could be used to test for biological activity. See Anderson et al., WO 99/09146, and Erikson et al., U.S. Patent 6,358,738 B1. Also, RNA-mediated interference ("RNAi") data has indicated that disruption of a PLK protein expression disrupts and/or inhibits oocyte and embryonic development in *Caenorhabdaitis elegans*. See Chase et al.. However, most of these methods tend to screen for pharmacologically active compounds rather than insecticidally active compounds. In addition, most of these methods involve mammalian or simple invertebrate organisms, such as, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster*, and *Mus musculus*, rather than more complex invertebrate organisms, such as organisms of the order of *Hemiptera*, in particular the invertebrate organism A*phis gossypii*. As a result, there is a need for a method for measuring the ability of a compound to act on a PLK protein of complex invertebrate organisms using an assay method such as inhibition of oocyte and embryonic development.

Summary of the Invention

One embodiment of the invention relates to nucleotide sequences that encode polypeptides with *Hemipteran* PLK activity. In preferred embodiments, nucleotide sequences that encode polypeptides with Aphid PLK activity. Such nucleotide sequences may be used to express amino acid sequences that are useful in the identification or development of compounds that affect PLK activity. Such compounds may be useful as pesticides or as pharmaceuticals. These nucleotide sequences that encode polypeptides with *Hemipteran* PLK activity, preferably Aphid PLK activity, including mutants and fragments thereof, which will be further described below, will also be referred to herein as "*nucleotide sequences of*

the invention". The polypeptides with *Hemipteran* PLK activity, preferably Aphid PLK activity, including mutants and fragments thereof, which will be further characterized below, will also be described as "proteins of the invention," "amino acid sequences of the invention, " or "polypeptide of the invention"

The present invention further relates to the use of the nucleotide sequences of the invention, preferably in the form of a suitable genetic construct as described below, in the transformation of host cells or host organisms, for example for the expression of the amino acid sequences of the invention and to such genetic constructs and host cells.

Another aspect of the invention relates to methods for the identification or development of compounds that can modulate and/or inhibit the biological activity of the amino acid sequences of the invention, in which one or more of the above mentioned nucleotide sequences, amino acid sequences, genetic constructs, host cells or host organisms are used. Such methods, which will usually be in the form of an assay or screen, will also be further described below.

A further aspect of the invention relates to compounds that can modulate the biological activity of, or that can otherwise interact with, an amino acid sequence of the invention, either *in vitro* or preferably (also) *in vivo*. The invention also relates to compositions that contain such compounds, and to the use of such compounds in the preparation of these compositions and the control of pests.

Definitions

Collectively, the nucleic acids of the present invention will be referred to herein as "*nucleic acids of the invention*". Also, where appropriate in the context of the further description of the invention below, the terms "*nucleotide sequence of the invention*" and "*nucleic acid of the invention*" may be considered essentially equivalent and essentially interchangeable.

Also, for the purposes of the present invention, a nucleic acid or amino acid sequence is considered to be "*(in) essentially isolated (form)*" – for example, from its native biological source - when it has been separated from at least one other nucleic acid molecule and/or sequence with which it is usually associated. Similarly, a protein or polypeptide of the invention is considered to be "*(in)*

essentially isolated (form)" – for example, from its native biological source - when it has been effectively separated from other polypeptide molecules with which it is normally associated. In particular, a nucleic acid or polypeptide of the invention is considered "essentially isolated" when it has been purified at least 2-fold, in particular at least 10-fold, more in particular at least 100-fold, and up to 1000-fold or more.

Detailed Description of the Invention

The present invention was established from the finding that the amino acid sequences of the invention can be used as "target(s)" in assays to identify chemical compounds and other factors (with the term "*target*" having its usual meaning in the art, provide for example the definition given in WO 98/06737) which interact with them *in vitro* or *in vivo*. Consequently, compounds or factors that have been identified as interacting with the amino acid sequences of the invention (e.g. by the methods as described herein below) may be useful as active agents in the agrochemical, veterinary or pharmaceutical fields.

In one embodiment, the invention relates to a nucleic acid, preferably in essentially isolated form, which nucleic acid comprises a nucleotide sequence of the invention, and in particular the nucleotide sequence of SEQ ID NO: 1 and SEQ ID NO: 3 and mutants and fragments thereof. The nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, preferably SEQ ID NO: 1, was derived or isolated from the *Aphis gossypii* organism, in the manner as further described in the Examples below. The present invention also relates to a polypeptides of the invention, and in particular those comprising SEQ ID NO: 2 or SEQ ID NO: 4, and mutants and fragments thereof. SEQ ID NO: 1 encodes SEQ ID NO: 2 and SEQ ID NO: 3 encodes SEQ ID NO: 4.

Generally, the nucleotide sequences of the invention, when in the form of a nucleic acid, may be DNA or RNA, and may be single stranded or double stranded. For example, the nucleotide sequences of the invention may be genomic DNA, cDNA or synthetic DNA (such as DNA with a codon usage that has been specifically adapted for expression in the intended host cell or host organism, which may for instance be designed using suitable computer programs such as the BackTranslate analysis tool in Vector NTI (InforMax, Inc., Bethesda, MD). Thus,

the nucleotide sequences of the invention may contain intron sequences, and also generally comprises different splice variants.

Yet another embodiment relates to a double stranded RNA molecule directed against a nucleotide sequence of the invention (one strand of which will usually comprise at least part of a nucleotide sequence of the invention). Such double stranded RNA molecules have particular utility in RNA interference studies of gene function (Zamore et al, Cell 101:25-33 (2000)). The invention also relates to genetic constructs that can be used to provide such double stranded RNA molecules (e.g. by suitable expression in a host cell or host organism, or for example in a bacterial strain such as *E.coli*). For such constructs, reference is made to Maniatis et al., *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Press, 1989).

In a broader sense, the term "*nucleotide sequence of the invention*" also comprises:

- parts or fragments of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, preferably SEQ ID NO: 1;
- (natural or synthetic) mutants, variants, alleles, analogs; orthologs (herein below collectively referred to as "*mutants*") of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, as further described below.
- parts or fragments of such (natural or synthetic) mutants;
- nucleotide fusions of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3 (or a part or fragment thereof) with at least one further nucleotide sequence;
- nucleotide fusions of (natural or synthetic) mutants (or a part or fragment thereof) with at least one further nucleotide sequence;

in which such mutants, parts, fragments or fusions are preferably as further described below. The invention also comprises different splice variants of the above nucleotide sequences.

In some embodiments, the nucleotide sequence of the invention is a fragment of a nucleic acid molecule that encodes a Hemipteran PLK. Preferably, a nucleotide sequence of the invention will have a length of at least 500 nucleotides, preferably at least 1,000 nucleotides, more preferably at least 1,200 nucleotides;

and up to a length of at most 3,500 nucleotides, preferably at most 3,000 nucleotides, more preferably at most, 2,240 nucleotides. Examples of parts or fragments of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, preferably SEQ ID NO: 1; or a part or fragment of a (natural or synthetic) mutant thereof include, but are not limited to, 5' or 3' truncated nucleotide sequences, or sequences with an introduced in frame startcodon or stopcodon. Also, two or more such parts or fragments of one or more nucleotide sequences of the invention may be suitably combined (e.g. ligated in frame) to provide a further nucleotide sequence of the invention.In some embodiments, such parts or fragments comprise at least one continuous stretch of at least 100 nucleotides, preferably at least 250 nucleotides, more preferably at least 500 nucleotides, even more preferably more than 1,000 nucleotides, of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, preferably SEQ ID NO: 1.

Also, it is expected that - based upon the disclosure herein - the skilled person will be able to identify, derive or isolate natural "mutants" (as mentioned above) of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, preferably SEQ ID NO: 1, from (other individuals of) the same species (for example from an individual of a different strain or line, including but not limited to mutant strains or lines). It is also expected that - based upon the disclosure herein - the skilled person will be able to provide or derive synthetic mutants (as defined hereinabove) of the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, preferably SEQ ID NO: 1. It is also expected that - based upon the disclosure herein - the skilled person will be able to provide or derive polypeptide having polypeptide sequences of SEQ ID NO: 2 by means of protein expression.

In one specific embodiment, the mutant is such that it encodes the nucleotide sequence of SEQ ID NO: 1 or a part or fragment thereof. In another specific embodiment, the mutant is such that it encodes the nucleotide sequence of SEQ ID NO: 3 or a part or fragment thereof

Preferably, any mutants as described herein will have one or more, and preferably all, of the structural characteristics or conserved features referred to below for the nucleotide sequences of SEQ ID NO: 1 and SEQ ID NO: 3.

In particular, any mutants, parts or fragments as described herein may be such that they at least encode the active or catalytic site of the corresponding amino acid sequence of the invention and a binding domain of the corresponding amino acid sequence of the invention.

Also, any mutants, parts or fragments as described herein will preferably have a degree of "sequence identity", at the nucleotide level, with the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, preferably SEQ ID NO: 1, of at least 75%, preferably at least 80%, more preferably at least 85%, and in particular more than 90%, and up to 95% or more.

Also, preferably, any mutants, parts or fragments of the nucleotide sequence of the invention will be such that they encode an amino acid sequence which has a degree of "sequence identity", at the amino acid level, with the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, preferably SEQ ID NO: 2, of at least 55%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95% or more, in which the percentage of "sequence identity" is calculated as described below.

For this purpose, the percentage of "sequence identity" between a given nucleotide sequence and the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3 may be calculated by dividing the number of nucleotides in the given nucleotide sequence that are identical to the nucleotide at the corresponding position in the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3 by the total number of nucleotides in the given nucleotide sequence and multiplying by 100%, in which each deletion, insertion, substitution or addition of a nucleotide - compared to the sequence of SEQ ID NO: 1 or SEQ ID NO: 3 - is considered as a difference at a single nucleotide position.

The preferred computer program for performing global sequence alignments and determining sequence identity is ClustalW (Higgins et al., *Nucleic Acids Research* 22:4673-4680 (1994)), which is publicly available for a variety of computer platforms. Preferably the parameters used with the ClustalW program for protein sequence alignments are ktuple = 1, diagonals = 5, windows = 5, gap = 3, score = PERCENTAGE, matrix = BLOSUM, open penalty = 10.0 and extension penalty = 0.05.

Also, in a preferred aspect, any mutants, parts or fragments as described herein will encode proteins or polypeptides having biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NO: 1 and SEQ ID NO: 3, preferably SEQ ID NO: 1, i.e. to a degree of at least 55%, preferably at least 75%, and up to 90%, as measured by standard assay techniques as described below.

Any mutants, parts or fragments as described herein are preferably such that they are capable of hybridizing with the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, preferably SEQ ID NO: 1, i.e. under conditions of "moderate stringency", and preferably under conditions of "high stringency". Such conditions will be clear to the skilled person, for example from the standard handbooks, such as Sambrook et al. and Ausubel et al., mentioned above, as well as in EP 0 967 284, EP 1 085 089 or WO 00/55318.

It is also within the scope of the invention to use a fusion of a nucleotide sequence of the invention (as described above) with one or more further nucleotide sequence(s), including but not limited to one or more coding sequences, non-coding sequences or regulatory sequences. Preferably, in such fusions, the one or more further nucleotide sequences are operably connected (as described below) to the nucleotide sequence of the invention (for example so that, when the further nucleotide sequence is a coding sequence, the nucleotide fusion encodes a protein fusion as described below).

In another embodiment, the invention relates to an antisense molecule against a nucleotide sequence of the invention.

A nucleic acid, preferably in essentially isolated form, can be used to express an amino acid sequence of the invention, for example, the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4.

On the basis of the above, and although the invention is not specifically limited to any specific explanation or mechanism, the nucleotide sequences of some embodiments of the invention encode proteins that have biological activity as a polo-like kinase from insects of the order *Hemiptera*, which include aphids, leafhoppers, whiteflies, scales and true bugs that have mouthparts adapted to piercing and sucking. The nucleic acids of the invention may also be in the form

of a genetic construct, again as further described below. Genetic constructs of the invention will generally comprise at least one nucleotide sequence of the invention, optionally linked to one or more elements of genetic constructs known per se, as described below. Such genetic constructs may be DNA or RNA, and are preferably double-stranded DNA. The constructs may also be in a form suitable for transformation of the intended host cell or host organism, in a form suitable for integration into the genomic DNA of the intended host cell or in a form suitable independent replication, maintenance and inheritance in the intended host organism. For instance, the genetic construct may be in the form of a vector, such as for example a plasmid, cosmid, a yeast artificial chromosome ("YAC"), a viral vector or transposon. In particular, the vector may be an expression vector, i.e. a vector that can provide for expression *in vitro* or *in vivo* (e.g. in a suitable host cell or host organism as described below). An expression vector comprising a nucleotide sequence of the invention is also referred to herein as a recombinant expression vector. These constructs will also be referred to herein as "*genetic constructs of the invention*".

In a preferred embodiment, such a construct a recombinant expression vector which will comprise:

a) the nucleotide sequence of the invention; operably connected to:

b) one or more regulatory elements, such as a promoter and optionally a suitable terminator;

and optionally also:

c) one or more further elements of genetic constructs known per se; in which the terms "*regulatory element*", "*promoter*", "*terminator*", "*further elements*" and "*operably connected*" have the meanings indicated herein below.

As the one or more "further elements" referred to above, the genetic construct(s) of the invention may generally contain one or more suitable regulatory elements (such as a suitable promoter(s), enhancer(s), or terminator(s)), 3'- or 5'-untranslated region(s) ("UTR") sequences, leader sequences, selection markers, expression markers or reporter genes, or elements that may facilitate or increase (the efficiency of) transformation or integration. These and other suitable elements for such genetic constructs will be clear to the skilled person, and may for instance

depend upon the type of construct used, the intended host cell or host organism; the manner in which the nucleotide sequences of the invention of interest are to be expressed (e.g. via constitutive, transient or inducible expression); and the transformation technique to be used.

Preferably, in the genetic constructs of the invention, the one or more further elements are "*operably linked*" to the nucleotide sequence(s) of the invention or to each other, by which is generally meant that they are in a functional relationship with each other. For instance, a promoter is considered "*operably linked*" to a coding sequence if said promoter is able to initiate or otherwise control or regulate the transcription or the expression of a coding sequence (in which said coding sequence should be understood as being "*under the control of*" said promoter)

Generally, when two nucleotide sequences are operably linked, they will be in the same orientation and usually also in the same reading frame. They will usually also be essentially contiguous, although this may also not be required.

Preferably, the optional further elements of the genetic construct(s) used in the invention are such that they are capable of providing their intended biological function in the intended host cell or host organism.

For instance, a promoter, enhancer or terminator should be "*operable*" in the intended host cell or host organism, by which is meant that (for example) said promoter should be capable of initiating or otherwise controlling or regulating the transcription or the expression of a nucleotide sequence - e.g. a coding sequence - to which it is operably linked (as defined above).

Such a promoter may be a constitutive promoter or an inducible promoter, and may also be such that it (only) provides for expression in a specific stage of development of the host cell or host organism, or such that it (only) provides for expression in a specific cell, tissue, organ or part of a multicellular host organism.

Some particularly preferred promoters include, but are not limited to, constitutive promoters, such as cytomegalovirus ("CMV"), Rous sarcoma virus ("RSV"), simian virus-40 ("SV40"), for example, pSVL SV40 Late Promoter Expression Vector (Pharmacia Biotech Inc., Piscataway, NJ), or herpes simplex virus ("HSV") for expression in mammalian cells or insect constitutive promoters

such as the immediate early baculovirus promoter described by Jarvis et al. (Methods in Molecular BiologyVol. 39 Baculovirus Expression Protocols, ed. C. Richardson., Hamana Press Inc., Totowa, NJ (1995)) available in pIE vectors from Novagen (Novagen, Inc. Madison, WI) or insect inducible promoters such as the *Drosophila metallothionein* promoter described by Bunch et al. (Nucleic Acids Research, Vol. 6, No. 3 1043-106, (1988)) available in vectors from Invitrogen (Invitrogen Corporation, Carlsbad, CA).

A selection marker should be such that it allows - i.e. under appropriate selection conditions - host cells or host organisms that have been (successfully) transformed with the nucleotide sequence of the invention to be distinguished from host cells or organisms that have not been (successfully) transformed. Some preferred, but non-limiting examples of such markers are genes that provide resistance against antibiotics (such as geneticin or G-418 (GIBCO- BRL, Grand Island, NY), kanamycin or ampicillin), genes that provide for temperature resistance, or genes that allow the host cell or host organism to be maintained in the absence of certain factors, compounds or (food) components in the medium that are essential for survival of the non-transformed cells or organisms.

A leader sequence should be such that - in the intended host cell or host organism - it allows for the desired post-translational modifications or such that it directs the transcribed mRNA to a desired part or organelle of a cell such as a signal peptide. A leader sequence may also allow for secretion of the expression product from said cell. As such, the leader sequence may be any pro-, pre-, or prepro-sequence operable in the host cell or host organism, including, but not limited to, picornavirus leaders, potyvirus leaders, a human immunoglobulin heavy-chain binding protein ("BiP"), a tobacco mosaic virus leader ("TMV"), and a maize chlorotic mottle virus leader ("MCMV").

An expression marker or reporter gene should be such that - in the host cell or host organism - it allows for detection of the expression of (a gene or nucleotide sequence present on) the genetic construct. An expression marker may optionally also allow for the localization of the expressed product, e.g. in a specific part or organelle of a cell or in (a) specific cell(s), tissue(s), organ(s) or part(s) of a multicellular organism. Such reporter genes may also be expressed as a protein

fusion with the amino acid sequence of the invention. Some preferred, but non-limiting examples include fluorescent proteins, such as GFP, antibody recognition proteins, for example, V5 epitope or poly Histidine available in vectors and antibodies supplied by Invitrogen, or purification affinity handles such as polyhistidine which allows for purification on nickel columns or dihydrofolate reductase which allows for purification on methotrexate column, or markers which allow for selection of cells expressing the gene such as the *E. coli* beta-galactosidase gene.

For some non-limiting examples of the promoters, selection markers, leader sequences, expression markers and further elements that may be present or used in the genetic constructs of the invention - such as terminators, transcriptional or translational enhancers or integration factors - reference is made to the general handbooks such as Sambrook et al. and Ausubel et al. mentioned above, to W.B. Wood et al., "*The nematode Caenorhabditis elegans*", Cold Spring Harbor Laboratory Press (1988) and D.L. Riddle et al., "*C. ELEGANS II*", Cold Spring Harbor Laboratory Press (1997), as well as to the examples that are given in WO 95/07463, WO 96/23810, WO 95/07463, WO 95/21191, WO 97/11094, WO 97/42320, WO 98/06737, WO 98/21355, U.S. Patent 6,207,410, U.S. Patent 5,693,492 and EP 1 085 089. Other examples will be clear to the skilled person.

Another embodiment of the invention relates to a host cell or host organism that has been transformed or contains with a nucleotide sequence, with a nucleic acid or with a genetic construct of the invention. The invention also relates to a host cell or host organism that expresses, or (at least) is capable of expressing (e.g. under suitable conditions), an amino acid sequence of the invention. Collectively, such host cells or host organisms will also be referred to herein as "*host cells or host organisms of the invention*".

The host cell may be any suitable (fungal, prokaryotic or eukaryotic) cell or cell line, for example:

- a bacterial strain, including but not limited to strains of *E. coli*, *Bacillus*, *Streptomyces* or *Pseudomonas*;
- a fungal cell, including but not limited to cells from species of *Aspergillus* or *Trichoderma*;

- a yeast cell, including but not limited to cells from species of *Kluyveromyces* or *Saccharomyces;*
- an amphibian cell or cell line, such as *Xenopus* oocytes.

In one specific embodiment, which may particularly useful when the nucleotide sequences of the invention are (to be) used in the discovery and development of insecticidal compounds, the host cell may be an insect-derived cell or cell line, such as:

- cells or cell lines derived from *Hemipteran*, including, but not limited to, *Spodoptera* SF9 and Sf21 cells and cells or cell lines derived from *Aphis*;
- cells or cell lines derived from *Drosophila*, such as Schneider and Kc cells; and
- cells or cell lines derived from a pest species of interest (as mentioned below), such as from *Aphis gossypii.*

The host cell may also be a mammalian cell or cell line, including but not limited to CHO- and BHK-cells and human cells or cell lines such as HEK, HeLa and COS.

The host organism may be any suitable multicellular (vertebrate or invertebrate) organism, including but not limited to:

- a nematode, including but not limited to nematodes from the genus *Caenorhabditis*, such as *C. elegans*,
- an insect, including but not limited to species of *Aphis*, *Drosophila, Heliothis,* or a specific pest species of interest (such as those mentioned above);
- other well known model organisms, such as zebrafish;
- a mammal such as a rat or mouse;

Other suitable host cells or host organisms will be clear to the skilled person, for example from the handbooks and patent applications mentioned above.

It should be noted that when a nucleotide sequence of the invention is expressed in a multicellular organism, it may be expressed throughout the entire organism, or only in one or more specific cells, tissues, organs or parts thereof, for example by expression under the control of a promoter that is specific for said cell(s), tissue(s), organ(s) or part(s).

The nucleotide sequence may also be expressed during only a specific stage of development or life cycle of the host cell or host organism, again for example by expression under the control of a promoter that is specific for said stage of development or life cycle. Also, as already mentioned above, said expression may be constitutive, transient or inducible.

Preferably, these host cells or host organisms are such that they express, or are (at least) capable of expressing (e.g. under suitable conditions), an amino acid sequence of the invention (and in case of a host organism: in at least one cell, part, tissue or organ thereof). The invention also includes further generations, progeny and offspring of the host cell or host organism of the invention, which may for instance be obtained by cell division or by sexual or asexual reproduction. The amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, preferably SEQ ID NO: 2, may be isolated from the species mentioned above, using any technique(s) for protein isolation and purification known to one skilled in the art. Alternatively, the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, preferably SEQ ID NO: 2, may be obtained by suitable expression of a suitable nucleotide sequence - such as the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3 or a suitable mutant thereof - in an appropriate host cell or host organism, as further described below.

In another aspect, the invention relates to a protein or polypeptide, preferably in (essentially) isolated form, said protein or polypeptide comprising an amino acid sequence of the invention (as defined above), in particular the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, more particularly preferred the amino acid sequence of SEQ ID NO 2.

In a broader sense, the term "*amino acid sequence of the invention*" also comprises:

- parts or fragments of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4;
- (natural or synthetic) mutants, variants, alleles, analogs, orthologs (herein below collectively referred to as "*analogs*") of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4;
- parts or fragments of such analogs;

- fusions of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 (or a part or fragment thereof) with at least one further amino acid residue or sequence;
- fusions of the amino acid sequence of an analog (or a part or fragment thereof) with at least one further amino acid residue or sequence;

in which such mutants, parts, fragments or fusions are preferably as further described below.

The term "*amino acid sequence of the invention*" also comprises "immature" forms of the above-mentioned amino acid sequences, such as a pre-, pro- or prepro-forms or fusions with suitable leader sequences. Also, the amino acid sequences of the invention may have been subjected to post-translational processing or be suitably glycosylated, depending upon the host cell or host organism used to express or produce said amino acid sequence; or may be otherwise modified (e.g. by chemical techniques known per se in the art).

Examples of parts or fragments of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, or a part or fragment of a (natural or synthetic) analog thereof mutant thereof include, but are not limited to, N- and C- truncated amino acid sequence. Also, two or more parts or fragments of one or more amino acid sequences of the invention may be suitably combined to provide an amino acid sequence of the invention.

Preferably, an amino acid sequence of the invention has a length of at least 100 amino acids, preferably at least 250 amino acids, more preferably at least 300 amino acids; and up to a length of at most 1,000 amino acids, preferably at most 750 amino acids, more preferably at most 600 amino acids.

Preferably, any such parts or fragments will be such that they comprise at least one continuous stretch of at least 5 amino acids, preferably at least 10 amino acids, more preferably at least 20 amino acids, even more preferably more than 30 amino acids, of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4.

In particular, any parts or fragments as described herein are such that they (at least) comprise the active or catalytic site of the corresponding amino acid sequence of the invention or a binding domain of the corresponding amino acid sequence of the invention. As will be clear to the skilled person, such parts or

fragments may find particular use in assay- and screening techniques (as generally described below) and (when said part or fragment is provided in crystalline form) in X-ray crystallography.

Also, it is expected that - based upon the disclosure herein - the skilled person will be able to identify, derive or isolate natural "analogs" (as mentioned above) of the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4. Such mutants could be derived from (other individuals of) the same species (for example from an individual of a different strain or line, including but not limited to mutant strains or lines); or from (individuals of) other species. For example, such analogs could be derived from the insect species mentioned above.

It is also expected that - based upon the disclosure herein - the skilled person will be able to provide or derive synthetic "analogs" (as mentioned above) of the amino sequence of SEQ ID NO: 2 or SEQ ID NO: 4, preferably SEQ ID NO: 2.

Preferably, any mutants as described herein will have one or more, and preferably all, of the structural characteristics or conserved features referred to below for the sequences of SEQ ID NO: 2 or SEQ ID NO: 4, preferably SEQ ID NO: 2.

Preferably, any analogs, parts or fragments as described herein will be such that they have a degree of "sequence identity", at the amino acid level, with the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4, preferably SEQ ID NO: 2, of at least 55%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95 % or more.

For this purpose, the percentage of "sequence identity" between a given amino acid sequence and the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 may be calculated by dividing the number of amino acid residues in the given amino acid sequence that are identical to the amino acid residue at the corresponding position in the amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 by the total number of amino acid residues in the given amino acid sequence and multiplying by 100%, in which each deletion, insertion, substitution or addition of an amino acid residue - compared to the sequence of SEQ ID NO: 2 or

SEQ ID NO: 4 - is considered as a difference at a single amino acid (position). As mentioned above, the preferred method of performing pairwise global sequence alignments fur such calculations is with the program ClustalW.

Also, preferably, any analogs, parts or fragments as described herein will have a biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NO: 2 and SEQ ID NO: 4, preferably SEQ ID NO: 2, i.e. to a degree of at least 10%, preferably at least 50% more preferably at least 75%, and up to 90%, as measured by standard assay techniques as described below.

It is also within the scope of the invention to use a fusion of an amino acid sequence of the invention (as described above) with one or more further amino acid sequences, for example to provide a protein fusion. Generally, such fusions may be obtained by suitable expression of a suitable nucleotide sequence of the invention - such as a suitable fusion of a nucleotide sequence of the invention with one or more further coding sequences - in an appropriate host cell or host organism, as further described below.

One particular embodiment, such fusions may comprise an amino acid sequence of the invention fused with a reporter protein such as glutathione S-transferase ("GST"), green fluorescent protein ("GFP"), luciferase or another fluorescent protein moiety. As will be clear to the skilled person, such fusions may find particular use in expression analysis and similar methodologies.

In another embodiment, the fusion partner may be an amino acid sequence or residue that may be used in purification of the expressed amino acid sequence, for example using affinity techniques directed against said sequence or residue. Thereafter, said sequence or residue may be removed (e.g. by chemical or enzymatical cleavage) to provide the nucleotide sequence of the invention (for this purpose, the sequence or residue may optionally be linked to the amino acid sequence of the invention via a cleavable linker sequence). Some preferred, but non-limiting examples of such residues are multiple histidine residues and glutatione residues.

In one preferred, but non-limiting aspect, any such fusion will have a biological activity that is essentially similar to the biological activity described

above for the sequences of SEQ ID NO: 2 and SEQ ID NO: 4, preferably SEQ ID NO: 2, i.e. to a degree of at least 10%, preferably at least 50 % more preferably at least 75%, and up to 90%, as measured by standard assay techniques as described below.

The nucleotide sequences and amino acid sequences of the invention may generally be characterized by the presence of one or more of the following structural characteristics or conserved features:

For the gene *Aphis gossypii*: SEQ ID NO: 1 is a cDNA sequence encompassing the open reading frame; SEQ ID NO: 2 is the protein encoded by SEQ ID NO: 1. SEQ ID NO: 3 is a partial cDNA sequence from an *Aphis gossypii* Expressed Sequencing Tag ("EST") library. SEQ ID NO: 4 is the polypeptide encoded by SEQ ID NO: 3. The *Aphis* PLK protein sequence is related to other polo-like kinases as set forth in the table below, where the relatedness values were determined using ClustalW and the parameters set forth above.

	Percentage Identity	
Sequence:	Full length	Kinase domain
Drosophila Polo [1]	52	64
Human PLK-1 [2]	52	68
C. elegans PLK-1 [3]	46	60
Aphis gossypii (SEQ ID NO: 2)	100	100

[1]Genebank Accession No. NP_524179, Llamazares, S et al., Genes Dev. 5 (12A), 2153-2165, 1991.
[2]Genebank Accession No. NP_005021, Holtrich, U. et al., Proc. Natl. Acad. Sci. U.S.A. 91 (5), 1736-1740, 1994.
[3]Genebank Accession No. AF080581, Chase, D. et al., Genesis 26, 26-41, 2000.

By analogy to other polo-like kinases, it is likely that the functional protein is monomeric. (*See, e.g.*, Hannan and Hall, In Comparative Molecular Neurobiology, Y. Pichon, 1993, Birkhuaser Verlag Basel Switzerland).

On the basis of the above, and although the invention is not specifically limited to any specific explanation or mechanism, the nucleotide sequences and amino acid sequences have (biological) activity as a serine and/or threonine kinase. In particular, the present invention has shown activity as a polo-like kinase from

insects of the order *Hemiptera*, which are aphids, leafhoppers, whiteflies, scales and true bugs that have mouthparts adapted to piercing and sucking.

As is known in the art, biological activity of this kind can be measured using standard assay techniques (see Erikson et al., U.S. Patent 6,358,738 B1), for example, by adenosine 5'-[γ-^{35}S]-triphosphate binding to the kinase; by fluorescent assays based on protein interactions such as fluorescence resonance energy transfer, time resolved fluorescence, fluorescence polarization or fluorometic or colorimetric reporter assays; or any technology suitable for assaying PLK proteins, for example, by measuring casein phosphorylation and/or PLK substrate phosphorylation. Such technology, as well as other well-known technology, can be adapted to methods of finding biologically active compounds that specifically effect the compositions of the present invention, particularly, the polypeptide of the invention. In preferred embodiments, the biological activity is the inhibition of *Hemipteran* PLK activity, preferably Aphid PLK activity. It is preferred that when performing test assays using test compounds, the test compounds be prepared as serial dilutions such that multiple test assays are performed with the same compound at different concentrations. In some embodiments, assays optionally employ positive controls such as compounds known to have a particular biological activity. In some embodiments, the positive control may be an antibody specific for the polypeptide of the invention. In some embodiments, assays optionally employ negative controls such as compounds known not to have a particular biological activity. These positive control assays and/or negative control assays may be run side by side with test assays to help establish and confirm that the compounds identified as having biological activity in the test assays are in fact biologically active. In addition to positive and negative control assays, additional control assays may be performed using other known PLKs from other species. By performing these types of assays, it may be established that the compound found to be active in the test assay is not active in assays using PLK from other species. This result would indicate that the compound has some degree of specificity. Kits can be provided to allow for such assays to be performed. Kits can include containers containing some or all of the reagents needed to perform the assay,

samples of a polypeptide of the invention, a nucleic acid of the invention, a genetic construct of the invention or a host cell of the invention.

Another embodiment of the invention relates to a nucleic acid probe that is capable of hybridizing with a nucleotide sequence of the invention under conditions of moderate stringency, preferably under conditions of high stringency, and in particular under stringent conditions (all as described above). Such nucleotide probes may for instance be used for detecting or isolating a nucleotide sequence of the invention or as a primer for amplifying a nucleotide sequence of the invention; all using techniques known per se, for which reference is again made to the general handbooks such as Sambrook et al. and Ausubel et al., mentioned above.

Preferably, when to be used for detecting or isolating another nucleotide sequence of the invention, such a nucleotide probe will usually have a length of between 15 and 100 nucleotides, and preferably between 20 and 80 nucleotides. When used as a primer for amplification, such a nucleotide probe will have a length of between 25 and 75 nucleotides, and preferably between 20 and 40 nucleotides.

Generally, such probes can be designed by the skilled person starting from a nucleotide sequence or amino acid sequence of the invention - and in particular the sequence of SEQ ID NO: 1 or SEQ ID NO: 2 - optionally using a suitable computer algorithm. Also, as will be clear to the skilled person, such probes may be degenerate probes. Probes and primers preferably have sequences that include unique sequences. A unique sequence is a sequence that is not found on other DNA molecules. The presence of unique sequences ensures that the probe or primer will not cross hybridize to identical sequences found on other genes. One skilled in the art can readily determine if a probe or primer contains unique sequences by first designing the probe or primer and then comparing the sequences thereof with sequences in databases of known nucleic acid sequences. Such comparisons are routinely performed by those skilled in the art.

In a further aspect, the invention relates to methods for preparing mutants and genetic constructs of the nucleotide sequences of the present invention.

Natural mutants of the nucleotide sequences of the present invention may be obtained in a manner essentially analogous to the method described in the Examples, or alternatively by:

- construction of a DNA library from the species of interest in an appropriate expression vector system, followed by direct expression of the mutant sequence;
- construction of a DNA library from the species of interest in an appropriate expression vector system, followed by screening of said library with a probe of the invention (as described below) or with a nucleotide sequence of the invention;
- isolation of mRNA that encodes the mutant sequence from the species of interest, followed by cDNA synthesis using reverse transcriptase;

or by any other suitable method(s) or technique(s) known per se, for which reference is for instance made to the standard handbooks, such as Sambrook et al., "Molecular Cloning: A Laboratory Manual" (2nd.ed.), Vols. 1-3, Cold Spring Harbor Laboratory Press (1989) and F. Ausubel et al., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987).

Techniques for generating such synthetic sequences of the nucleotide sequences of the present invention will be clear to the skilled person and may for instance include, but are not limited to, automated DNA synthesis; site-directed mutagenesis; combining two or more parts of one or more naturally occurring sequences, introduction of mutations that lead to the expression of a truncated expression product; introduction of one or more restriction sites (e.g. to create cassettes or regions that may easily be digested or ligated using suitable restriction enzymes), and the introduction of mutations by means of a PCR reaction using one or more "mismatched" primers, using for example a sequence of other techniques will be clear to the skilled person, and reference is again made to the standard handbooks, such naturally occurring polo-like kinase as a template. These and as Sambrook et al. and Ausubel et al., mentioned above.

The genetic constructs of the invention may generally be provided by suitably linking the nucleotide sequence(s) of the invention to the one or more further elements described above, for example using the techniques described in

the general handbooks such as Sambrook et al. and Ausubel et al., mentioned above.

Often, the genetic constructs of the invention will be obtained by inserting a nucleotide sequence of the invention in a suitable (expression) vector known per se. Some preferred, but non-limiting examples of suitable expression vectors include:

- vectors for expression in mammalian cells: pSVL SV40 (Pharmacia), pMAMneo (Clontech), pcDNA3 (Invitrogen), pMC1neo (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593), pBPV-1 (8-2) (ATCC 37110), pdBPV-MMTneo (342-12) (ATCC 37224), pRSVgpt (ATCC37199), pRSVneo (ATCC37198), pSV2-dhfr (ATCC 37146), pUCTag (ATCC 37460) and 1ZD35 (ATCC 37565);
- vectors for expression in bacterials cells: pET vectors (Novagen), pGEX vectors (Pharmacia) and pQE vectors (Qiagen);
- vectors for expression in yeast or other fungal cells: pYES2 (Invitrogen) and Pichia expression vectors (Invitrogen);
- vectors for expression in insect cells: pBlueBacII (Invitrogen), pEI1 (Novagen), pMT/V5His (Invitrogen).

In a further aspect, the invention relates to methods for transforming a host cell or a host organism with a nucleotide sequence, with a nucleic acid or with a genetic construct of the invention. The invention also relates to the use of a nucleotide sequence, of a nucleic acid or of a genetic construct of the invention transforming a host cell or a host organism.

According to one specific embodiment, the expression of a nucleotide sequence of the invention in a host cell or host organism may be reduced, compared to the original (e.g. native) host cell or host organism. This may for instance be achieved in a transient manner using antisense or RNA-interference techniques well known in the art, or in a constitutive manner using random, site specific or chemical mutagenesis of the nucleotide sequence of the invention.

Suitable transformation techniques will be clear to the skilled person and may depend on the intended host cell or host organism and the genetic construct to be used. Some preferred, but non-limiting examples of suitable techniques include

ballistic transformation, (micro-)injection, transfection (e.g. using suitable transposons), electroporation and lipofection. For these and other suitable techniques, reference is again made to the handbooks and patent applications mentioned above.

After transformation, a step for detecting and selecting those host cells or host organisms that have been successfully transformed with the nucleotide sequence or genetic construct of the invention may be performed. This may for instance be a selection step based on a selectable marker present in the genetic construct of the invention or a step involving the detection of the amino acid sequence of the invention, e.g. using specific antibodies.

The transformed host cell (which may be in the form or a stable cell line) or host organisms (which may be in the form of a stable mutant line or strain) form further aspects of the present invention.

In yet another aspect, the invention relates to methods for producing an amino acid sequence of the invention.

To produce or obtain expression of the amino acid sequences of the invention, a transformed host cell or transformed host organism may generally be kept, maintained or cultured under conditions such that the (desired) amino acid sequence of the invention is expressed or produced. Suitable conditions will be clear to the skilled person and will usually depend upon the host cell or host organism used, as well as on the regulatory elements that control the expression of the (relevant) nucleotide sequence of the invention. Again, reference is made to the handbooks and patent applications mentioned above in the paragraphs on the genetic constructs of the invention.

Generally, suitable conditions may include the use of a suitable medium, the presence of a suitable source of food or suitable nutrients, the use of a suitable temperature, and optionally the presence of a suitable inducing factor or compound (e.g. when the nucleotide sequences of the invention are under the control of an inducible promoter); all of which may be selected by the skilled person. Again, under such conditions, the amino acid sequences of the invention may be expressed in a constitutive manner, in a transient manner, or only when suitably induced.

It will also be clear to the skilled person that the amino acid sequence of the invention may (first) be generated in an immature form (as mentioned above), which may then be subjected to post-translational modification, depending on the host cell or host organism used. Also, the amino acid sequence of the invention may be glycosylated, again depending on the host cell or host organism used.

The amino acid sequences of the invention may then be isolated from the host cell or host organism or from the medium in which said host cell or host organism was cultivated, using protein isolation and purification techniques known per se, such as (preparative) chromatography and electrophoresis techniques, differential precipitation techniques, affinity techniques (e.g. using a specific, cleavable amino acid sequence fused with the amino acid sequence of the invention) and preparative immunological techniques (i.e. using antibodies against the amino acid sequence to be isolated).

In one embodiment, the amino acid sequence thus obtained may also be used to generate antibodies specifically against said sequence or an antigenic part or epitope thereof.

In one embodiment, the present invention relates to antibodies, for example monoclonal and polyclonal antibodies, that are generated specifically against amino acid sequences of the present invention, preferably SEQ ID NO: 2, or an analog, variant, allele, ortholog, part, fragment or epitope thereof or SEQ ID NO: 4, or an analog, variant, allele, ortholog, part, fragment or epitope thereof. In preferred embodiments, the antibody does not cross reactive with other PLK polypeptides.

Such antibodies, which form a further aspect of the invention, may be generated in a manner known per se, for example as described in GB-A-2 357 768, USA 5,693,492, WO 95/32734, WO 96/23882, WO 98/02456, WO 98/41633 and WO 98/49306. Often, but not exclusively, such methods will involve as immunizing a immunocompetent host with the pertinent amino acid sequence of the invention or an immungenic part thereof (such as a specific epitope), in amount(s) and according to a regimen such that antibodies against said amino acid sequence are raised, and than harvesting the antibodies thus generated, e.g. from blood or serum derived from said host.

For instance, polyclonal antibodies can be obtained by immunizing a suitable host such as a goat, rabbit, sheep, rat, pig or mouse with (an epitope of) an amino acid sequence of the invention, optionally with the use of an immunogenic carrier (such as bovine serum albumin or keyhole limpet hemocyanin) or an adjuvant such as Freund's, saponin, aluminium hydroxide or a similar mineral gel, or keyhole limpet hemocyanin or a similar surface active substance. After a suitable immune response has been raised (usually within 1-7 days), the antibodies can be isolated from blood or serum taken from the immunized animal in a manner known per se, which optionally may involve a step of screening for an antibody with desired properties (i.e. specificity) using known immunoassay techniques, for which reference is again made to for instance WO 96/23882.

Monoclonal antibodies may for example be produced using continuous cell lines in culture, including hybridoma-based and similar techniques, again essentially as described in the above cited references. Accordingly, cells and cell lines that produce monoclonal antibodies against an amino acid sequence of the invention form a further aspect of the invention, as do methods for producing antibodies against amino acid sequences of the invention, which methods may generally involve cultivating such a cell and isolating the antibodies from the culture or medium, again using techniques known per se.

Also, Fab-fragments against the amino acid sequences of the invention (such as F (ab)$_2$, Fab' and Fab fragments) may be obtained by digestion of an antibody with pepsin or another protease, reducing disulfide-linkages and treatment with papain and a reducing agent, respectively. Fab-expression libraries may for instance be obtained by the method of Huse et al., 1989, Science 245:1275-1281.

In another embodiment, the amino acid sequence of the invention, or a host cell or host organism that expresses such an amino acid sequence, may also be used to identify or develop compounds or other factors that can modulate the (biological) activity of, or that can otherwise interact with, the amino acid sequences of the invention, and such uses form further aspects of the invention. As will be clear to the skilled person, in this context, the amino acid sequence of the invention will serve as a target for interaction with such a compound or factor.

In this context, the terms "*modulate*", "*modulation,* "*modulator*" and "*target*" will have their usual meaning in the art, for which reference is *inter alia* made to the definitions given in WO 98/06737. Generally, a modulator is a compound or factor that can enhance, inhibit or reduce or otherwise alter, influence or affect (collectively referred to as "*modulation*") a functional property of a biological activity or process (for example, the biological activity of an amino acid sequence of the invention). As noted above, it is known in the art that PLK activity can be measured using standard assay techniques (see Erikson et al., U.S. Patent 6,358,738 B1), for example, by adenosine 5'-[γ-^{35}S]-triphosphate binding to the kinase; by fluorescent assays based on protein interactions such as fluorescence resonance energy transfer, time resolved fluorescence, fluorescence polarization or fluorometic or colorimetric reporter assays; or any technology suitable for assaying PLK proteins, for example, by measuring casein phosphorylation and/or PLK substrate phosphorylation. Such technology, as well as other well-known technology, can be adapted to methods of finding biologically active compounds that specifically effect the compositions of the present invention, particularly, the polypeptide of the invention. In preferred embodiments, the biological activity is the inhibition of *Hemipteran* PLK activity, preferably Aphid PLK activity. It is preferred that when performing test assays using test compounds, the test compounds be prepared as serial dilutions such that multiple test assays are performed with the same compound at different concentrations. In some embodiments, assays optionally employ positive controls such as compounds known to have a particular biological activity. In some embodiments, the positive control may be an antibody specific for the polypeptide of the invention. In some embodiments, assays optionally employ negative controls such as compounds known not to have a particular biological activity. These positive control assays and/or negative control assays may be run side by side with test assays to help establish and confirm that the compounds identified as having biological activity in the test assays are in fact biologically active. In addition to positive and negative control assays, additional control assays may be performed using other known PLKs from other species. By performing these types of assays, it may be

established that the compound found to be active in the test assay is not active in assays using PLK from other species. This result would indicate that the compound has some degree of specificity. Kits can be provided to allow for such assays to be performed. Kits can include containers containing some or all of the reagents needed to perform the assay, samples of a polypeptide of the invention, a nucleic acid of the invention, a genetic construct of the invention or a host cell of the invention.

In this context, the amino acid sequence of the invention may serve as a target for modulation *in vitro* (e.g. as part of an assay or screen) or for modulation *in vivo* (e.g. for modulation by a compound or factor that is known to modulate the target, which compound or factor may for example be used as an active compound for agrochemical, veterinary or pharmaceutical use).

For example, the amino acid sequences, host cells or host organisms of the invention may be used as part of an assay or screen that may be used to identify or develop modulators of the amino acid sequence of the invention, such as a primary screen (e.g. a screen used to identify modulators of the target from a set or library of test chemicals with unknown activity with respect to the target) or a secondary assay (e.g. an assay used for validating hits from a primary screen or used in optimizing hit molecules, e.g. as part of hits-to-leads chemistry).

For instance, such an assay or screen may be configured as an *in vitro* assay or screen. Suitable techniques for such *in vitro* screening will be clear to the skilled person, and are for example described in Eldefrawi et al., (1987). FASEB J., Vol.1, pages 262-271 and Rauh et al., (1990), Trends in Pharmacol. Sci., vol.11, pages 325-329.

Such an assay or screen may also be configured as a cell-based assay or screen, in which a host cell of the invention is contacted with or exposed to a test chemical, upon which at least one biological response by the host cell is measured.

Also, such an assay or screen may also be configured as an whole animal screen, in which a host organism of the invention is contacted with or exposed to a test chemical, upon which at least one biological response (such as a phenotypical, behavioral or physiological change, including but not limited to paralysis or death) by the host organism is measured.

Thus, generally, the assays and screens described above will comprise at least one step in which the test chemical is contacted with the target (or with a host cell or host organism that expresses the target), and in particular in such a way that a signal is generated that is representative for the modulation of the target by the test chemical. In a further step, said signal may then be detected.

Accordingly, in one aspect, the invention relates to a method for generating a signal that is representative for the interaction of an amino acid sequence of the invention with a test chemical, said method at least comprising the steps of:

a) contacting the amino acid sequence of the invention, or a host cell or host organism containing or expressing an amino acid sequence, with said test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said amino acid sequence; and optionally

b) detecting the signal that may thus be generated.

In another aspect, the invention relates to a method for identifying modulators and/or inhibitors of an amino acid sequence of the invention (e.g. from a set or library of test chemicals), said method at least comprising the steps of:

a) contacting the amino acid sequence of the invention, or a host cell or host organism containing or expressing an amino acid sequence, with a test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said the target; and optionally

b) detecting the signal that may thus be generated, said signal identifying the modulator and/or inhibitor of said amino acid sequence.

Accordingly, the present invention provides methods of identifying a modulator and/or inhibitor of a *Hemipteran* polo-like kinase protein activity. In preferred embodiments, the *Hemipteran* polo-like kinase protein used in the methods has an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, a mutant thereof, a fragment thereof, SEQ ID NO: 4, a mutant thereof, and a fragment thereof, preferably SEQ ID NO: 2. In some embodiments, the nucleic acid sequence that encodes the *Hemipteran* polo-like kinase is SEQ ID

NO: 1 a fragment thereof, SEQ ID NO: 3, or a fragment thereof, preferably SEQ ID NO: 1.

A test chemical may be part of a set or library of compounds, which may be a diverse set or library or a focussed set or library, as will be clear to the skilled person. The libraries that may be used for such screening can be prepared using combinatorial chemical processes known in the art or conventional means for chemical synthesis.

The assays and screens of the invention may be carried out at medium throughput to high throughput, for example in an automated fashion using suitable robotics. In particular, in this embodiment, the method of the invention may be carried out by contacting the target with the test compound in a well of a multi-well plate, such as a standard 24, 96, 384, 1536 or 3456 well plate.

Usually, in a screen or assay of the invention, for each measurement, the target or host cell or host organism will be contacted with only a single test compound. However, it is also within the scope of the invention to contact the target with two or more test compounds - either simultaneously or sequentially - for example to determine whether said combination provides a synergistic effect.

Once a test chemical has been identified as a modulator and/or inhibitor for an amino acid sequence of the invention (e.g. by means of a screen or assay as described hereinabove), it may be used per se as a modulator and/or inhibitor of the relevant amino acid sequence of the invention, preferably, an amino acid sequence of SEQ ID NO: 2, a mutant thereof, a fragment thereof, SEQ ID NO: 4, a mutant thereof, and a fragment thereof, more preferably SEQ ID NO: 2 (e.g. as an active substance for agrochemical, veterinary or pharmaceutical use), or it may optionally be further optimized for final use, e.g. to improve properties such as solubility, adsorption, bio-availability, toxicity, stability, persistence, environmental impact, etc.. It will be clear to the skilled person that the nucleotide sequences, preferably SEQ ID NO: 1 or SEQ ID NO: 3, more preferably SEQ ID NO: 1, amino acid sequences, host cells or host organisms and methods of the invention may find further use in such optimization methodology, for example as (part of) secondary assays.

The invention is not particularly limited to any specific manner or mechanism in or via which the modulator and/or inhibitor (e.g. the test chemical, compound or factor) modulates, inhibits, or interacts with, the target (*in vivo* or *in vitro*). For example, the modulator and/or inhibitor may be an agonist, an antagonist, an inverse agonist, a partial agonist, a competitive inhibitor, a non-competitive inhibitor, a cofactor, an allosteric inhibitor or other allosteric factor for the target, or may be a compound or factor that enhances or reduces binding of target to another biological component associated with its (biological) activity, such as another protein or polypeptide, a receptor, or a part of organelle of a cell. As such, the modulator and/or inhibitor may bind with the target (at the active site, at an allosteric site, at a binding domain or at another site on the target, e.g. covalently or via hydrogen bonding), block and/or inhibit the active site of the target (in a reversible, irreversible or competitive manner), block and/or inhibit a binding domain of the target (in a reversible, irreversible or competitive manner), or influence or change the conformation of the target.

As such, the test chemical, modulator and/or inhibitor may for instance be:

- an analog of a known substrate of the target;
- an oligopeptide, e.g. comprising between 2 and 20, preferably between 3 and 15 amino acid residues;
- an antisense or double stranded RNA molecule;
- a protein, polypeptide;
- a cofactor or an analog of a cofactor.

The test chemical, modulator and/or inhibitor may also be a reference compound or factor, which may be a compound that is known to modulate, inhibit or otherwise interact with the target (e.g. a known substrate or inhibitor for the target) or a compound or factor that is generally known to modulate, inhibit or otherwise interact with other members from the general class to which the target belongs (e.g. a known substrate or inhibitor of said class).

Preferably, however, the test chemical, modulator and/or inhibitor is a small molecule, by which is meant a molecular entity with a molecular weight of less than 1500, preferably less than 1000. This may for example be an organic, inorganic or organometallic molecule, which may also be in the form or a suitable

salt, such as a water-soluble salt. The term "small molecule" also covers complexes, chelates and similar molecular entities, as long as their (total) molecular weight is in the range indicated above.

As already mentioned above, the compounds or factors that have been identified or developed as modulators and/or inhibitors of the amino acid sequences of the invention, preferably, an amino acid sequence of SEQ ID NO: 2, a mutant thereof, a fragment thereof, SEQ ID NO: 4, a mutant thereof, and a fragment thereof, more preferably SEQ ID NO: 2, (and precursors for such compounds) may be useful as active substances in the agrochemical, veterinary or pharmaceutical fields, for example in the preparation of agrochemical, veterinary or pharmaceutical compositions, and both such modulators as well as compositions containing them further aspects of the invention.

For example, in the agrochemical field, the modulators and/or inhibitors of the invention may be used as an insecticide, nematicide, molluscide, helminticide, acaricide or other types of pesticides or biocides, e.g. to prevent or control (infestations with) harmful organisms, both as contact agents and as systemic agents. As such, the modulators and/or inhibitors may for example be used as a crop protection agent, as a pesticide for household use, or as an agent to prevent or treat damage caused by harmful organisms (e.g. for the protection of seed, wood or stored crops or fruits). Preferably, the modulators and/or inhibitors of the invention are used as insecticides.

For any such application, one or more modulators and/or inhibitors of the invention may be suitably combined with one or more agronomically acceptable carriers, adjuvants or diluents - and optionally also with one or more further compounds known per se with activity as (for example) a plant protection agent (to broaden the spectrum of action and optionally to provide a synergistic effect), herbicide, fertilizer or plant growth regulator - to provide a formulation suitable for the intended final use. Such a formulation may for example be in the form of a solution, emulsion, dispersion, concentrate, aerosol, spray, powder, flowable, dust, granule, pellet, fumigation candle, bait or other suitable solid, semi-solid or liquid formulation, and may optionally also contain suitable solvents, emulsifiers, stabilizers, surfactants, antifoam agents, wetting agents, spreading agents, sticking

agents, attractants or (for a bait) food components. Reference is made to the standard manuals, such as "Pesticidal Formulation Research", ACS-publications (1969) and "Pesticide Formulations", Wade van Valkenburg Ed, Marcel Dekker publications (1973).

Such compositions may generally contain one or more modulators and/or inhibitors of the invention in a suitable amount, which generally may be between 0.1 and 99 %, and in particular between 10 and 50 %, by weight of the total composition.

The modulators and/or inhibitors and compositions of the invention may be particularly useful as insecticides, for example to combat or control undesired or harmful insects (both adult and immature forms, such as larvae) from following orders:

- *Coleoptera*, such as *Pissodes strobi*, *Diabrotica undecimpunctata howardi*, and *Leptinotarsa decemlineata*;
- *Diptera*, such as *Rhagoletis pomonella*, *Mayetiola destructor*, and *Liriomyza huidobrensis*;
- *Hymenoptera*, such as *Neodiprion taedae tsugae*, *Camponotus pennsylvanicus*, and *Solenopsis wagneri*;
- *Hemiptera*, such as *Pseudatomoscelis seriatus*, *Lygus lineolaris* (Palisot de Beauvois), *Acrosternum hilare*, and *Aphis gossypii*
- *Homoptera*; and
- *Lepidoptera* such as *Heliothis virescens*.

When used to control harmful or undesired organisms, these organisms may be directly contacted with the modulators, inhibitors, or compositions of the invention in an amount suitable to control (e.g. kill or paralyze) the organism. This amount may be readily determined by the skilled person (e.g. by testing the compound on the species to be controlled) and will usually be in the region of between particular between 10 and 500 g/ha, in particular between 100 and 250 g/ha.

The modulators, inhibitors, or compositions of the invention may also be applied systemically (e.g. to the habitat of the organism to be controlled or to the soil), and may also be applied to the plant, seed, fruit etc. to be protected, again in

suitable amounts, which can be determined by the skilled person. The modulators and/or inhibitors of the invention may also be incorporated - e.g. as additives - in other compositions known per se, for example to replace other pesticidal compounds normally used in such compositions.

In one specific embodiment, the modulators and/or inhibitors and compositions of the invention may be used in the fields of agrochemical, veterinary or human health to prevent or treat infection or damage or discomfort caused by parasitic organisms, and in particular by parasitic arthropods, nematodes and helminths such as:

- ectoparasitic arthropods such as ticks, mites, fleas, lice, stable flies, horn flies, blowflies and other biting or sucking ectoparasites;
- endoparasites organisms such as helminths;

and also to prevent or treat diseases that are caused or transferred by such parasites. For such purposes, the modulators and/or inhibitors of the invention may for example be formulated as a tablet, an oral solution or emulsion, an injectable solution or emulsion, a lotion, an aerosol, a spray, a powder, a dip or a concentrate.

In the fields of animal and human health, the modulators, inhibitors, and compositions of the invention may also be used for the prevention or treatment of diseases or disorders in which the amino acid sequence of the invention may be involved as a target. For this purpose, the modulators and/or inhibitors of the invention may be formulated with one or more additives, carriers or diluents acceptable for pharmaceutical or veterinary use, which will be clear to the skilled person.

Thus, in a further aspect, the invention relates to the use of a modulator and/or inhibitor of the invention in the preparation of a composition for agrochemical, veterinary or pharmaceutical use, as described hereinabove. The invention relates to the use of the modulators, inhibitors and compositions of the invention in controlling harmful organisms and in preventing infestation or damage caused by harmful organisms, again as described above.

The invention will now be further illustrated by means of the following non-limiting Examples.

Examples

Example 1 – Cotton Aphid ("CA") Polo-like Kinase Sequence Identifications

CA polyA RNA isolation. A 0.1% solution of diethyl pyrocarbonate ("DEPC" available from Aldrich Chemical Co., Inc. Milwaukee, WI) in water was incubated at 37°C for about 16 hours and then autoclaved for 60 minutes. All glassware was baked for four hours at 250°C and all bottle caps were soaked in the 0.1% DEPC solution prior to use. The microprobe of a Braun homogenizer (available from B. Braun Biotech International, Allentown, PA) was soaked in 50 mls of 100% ethanol ("EtOH", available from J.T. Baker Inc., Phillipsburg, NJ) and then run in 25 mls RNAzolB (a guanidinium hydrochloride preparation from CINNA-BIOTECX Labs, Inc., Houston, TX). Cotton aphids were collected from cotton plants and placed on ice in tared centrifuge tubes. After harvesting approximately 1.2 grams of material, the cotton aphids were frozen at -70°C until use. The aphids (1.0 gram) were then weighed out into two 0.5 gram aliquots. The aliquots were taken up in four mls of TRIZOL reagent (available from Invitrogen Corp., Carlsbad, CA). The resulting solution was homogenized at full speed and ambient temperature for thirty seconds. After this time, two mls of chloroform (available from J.T. Baker, Inc.) were added and the resulting mixture was centrifuged at 12000g and ambient temperature for fifteen minutes in an SS34 rotor (available from Sorvall Products, L.P, Asheville, NC). Upon completion of this period, the top aqueous layer was separate and placed in a fresh 15ml conical tube and an equal volume of isopropanol (available from J.T. Baker, Inc.) was added. The resulting mixture was centrifuged at 12000g and ambient a temperature for ten minutes in the SS34 rotor. The resulting pellet was then washed with 10 mls of an aqueous 75% EtOH solution and centrifuged at 7500g and ambient temperature for five minutes in the SS34 rotor. The supernatant was poured off and the final pellet was allowed to dry at ambient temperature for five minutes. After this time, the pellet was taken up in 1 ml of the 0.1% DEPC solution disclosed above. The concentration of total RNA was measured by UV spectrometry (GeneQuant from Amershan Biosciences, Piscataway, NJ). Five hundred µgs of total RNA were used to isolate polyA mRNA following the protocol of the Micro-FastTrack™ 2.0 kit (a available from Invitrogen Corp.). The

concentration of polyA mRNA was measured by UV spectrometry. The aphid polyA mRNA solution was stored at -70°C for future use.

Synthesis of 5'-RACE ready cDNA: The aphid polyA mRNA, 250 ng, was treated with a calf intestinal phosphatase (contained in a Invitrogen GeneRacer™ kit available from Invitrogen Corp.) according to the manufacturer's instructions to remove the 5' phosphates. The dephosphorylated RNA was then treated with a tobacco acid pyrophosphatase (contained in the Invitrogen GeneRacer™ kit) according to the manufacturer's instructions to remove the 5' cap structure from the intact, full-length mRNA. The GeneRacer™ RNA Oligo primer (contained in the Invitrogen GeneRacer™ kit) was then ligated to the 5' end of the mRNA using T4 RNA liganse (contained in the Invitrogen GeneRacer™ kit) according to the manufacturer's instructions. The ligated mRNA was then reverse-transcribed with random primers using cloned Avian Myeloblastosis Virus ("AMV") reverse transcriptase ("RT") (available from Invitrogen Corp.). The reactions were placed on a Geneamp 9700 thermal cycler (available from Perkin-Elmer-ABI, Foster City, CA) and held for 60 minutes at 42°C. After this time, the AMV RT was inactivated by heating at 85°C for 15 minutes followed by 2 minutes at 4°C. The resulting 5'-RACE ready cDNA was used for PCR amplification or stored at -20°C for future use.

PCR amplification. A 50 µl cDNA reaction was made utilizing buffers, dNTPs, and platinum Taq DNA polymerase high fidelity supplied in an Invitrogen GeneRacer™ kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. The PCR products were characterized by agarose gel electrophoresis. The bands were excised from NuSieve gels (FMC Corp., Philadelphia, PA) and purified using a Qiagen gel extraction Kit (QIAGEN Corp., Chatsworth, CA) according to the manufacture's instructions.

Primer synthesis and design. Oligonucleotides were synthesized by SeqWright Inc. (Houston, TX) and provided as lyophilized pellets which were dissolved in distilled water prior to use. PCR primers and probes were designed and annealing temperatures estimated using the VectorNTI suite 6.0 software from InforMax, Inc. (Bethesda, MD).

Subcloning and sequencing. Purified PCR fragments were subcloned into pCR2.1-TOPO vector using a TOPO TA Cloning kit (available from Invitrogen Corp.) according to the manufacturer's instructions. The resulting clones were sequenced by SeqWright Inc. and then analyzed using VectorNTI suite 6.0 software.

Full length cDNA assembly. A fragment containing the 5' end of the PLK gene was generated by digesting the 5' clone with the EcoRI site from the pCR2.1-TOPO cloning vector and an internal PvuII site approximately 1100bp downstream of the ATG. The fragment was ligated to a PvuII/Not I digested fragment from SEQ ID NO: 3. The resulting fragment was identified as SEQ ID NO: 1. The full length EcoRI/NotI fragment was then subcloned into an EcoRI/NotI digested pSK(-) (available from Stratagene, La Jolla, CA).

Primers. The primers utilized were as follows:

Primer	Sequence	Translation	Orientation
1	CGACTGGAGCACGAGGACACTGA (SEQ ID NO:5) 5'- INVITROGEN GENERACER™ PRIMER		forward
2	AATGAGAGGTTGAGCAGCTGG (SEQ ID NO:6)	PAAQPLI (SEQ ID NO:7)	reverse
3	AAAGCTCCTCTGAACGTCGCAAC (SEQ ID NO:8)		forward
4	CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAG UAGAAA (SEQ ID NO:9) GENERACER™ RNA OLIGO		forward

Aphis gossypii- Sequence Amplifications

SEQ ID NO: 3 was derived from FMC's proprietary *Aphis gossypii* Expressed Sequencing Tag ("EST") library. The *Aphis gossypii* sequence was then extended upstream by reverse transcribing the mRNA with random primers, followed by 5' RACE utilizing the Invitrogen GeneRacer™ kit and platinum Taq DNA polymerase high fidelity. Touchdown PCR with primers 1 and 2 generated sequence information from nucleotides 1 to 1386 in SEQ ID NO: 1. The amplimers from the above reactions were isolated, cloned, and sequenced. The

final 5'-end PCR fragment was generated by primers 2 and 3. This PCR fragment was then subcloned into a pCR2.1-TOPO vector. Subsequently, the 5' clone with EcoRI site from the pCR2.1-TOPO vector was digested with an internal PvuII site at nucliotide 1370 in SEQ ID NO: 1 to generate a fragment containing the 5' end of the *Aphis* PLK gene. This fragment was then ligated to a PvuII/Not I digested fragment from EST clone for SEQ ID NO: 3. The full length EcoRI/NotI fragment was then cloned into an EcoRI/NotI digested pSK (-).